

DOCUMENT RESUME

ED 172 906

PS 010 638

AUTHOR Taub, Herman P.
TITLE An Overview of a Project to Develop Head Start Profiles of Program Effects on Children.
INSTITUTION Mediarx Associates, Inc., Westport, Conn.
SPONS AGENCY Administration for Children, Youth, and Families (DHEW), Washington, D.C.
PUB DATE Apr 79
NOTE 29p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, California, April 8-12, 1979)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Cognitive Development; *Compensatory Education; Emotional Development; Individual Development; Interpersonal Competence; *Measurement Goals; Physical Development; Preschool Education; *Preschool Programs; *Profile Evaluation; *Program Effectiveness; Program Evaluation; Social Development
IDENTIFIERS *Development of Head Start Profiles Program Effects; *Project Head Start National Evaluation Program

ABSTRACT

This paper provides an overview of a project (now in its formative stages) to develop Head Start measures of program effects on children 3 to 7 years of age. The program approach is distinctive in emphasizing evaluation of all aspects of children's development that may affect their overall competence in reflecting the concerns, views, and values of a broad range of persons, and in recommending that the proposed measures be processed differently with children from different backgrounds in order to generate a plurality of scales of development. Specific descriptions of desired characteristics of children were collected from parents, public school teachers, Head Start staff members and child development scholars. Also, research literature was viewed, and special studies were commissioned. More than 1700 characteristics were identified and aggregated to form six domains: (1) health and physical development, (2) cognitive development, (3) social-emotional development, (4) aesthetic development, (5) moral and ethical development and (6) life skills. These domains include 26 dimensions of development. All domains except (4) and (5) are recommended for inclusion as evaluation measures. It is proposed that these measures be used to assess children's developmental status and that program effects be evaluated in terms of the relative change in children's development resulting from the program. Standards of significant change would be set in terms of the maximum change affected by the consistently most effective programs in a particular dimension for particular groups of children. (SS)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED172906

PS 010638

AN OVERVIEW OF A PROJECT TO DEVELOP HEAD START PROFILES
OF PROGRAM EFFECTS ON CHILDREN*

Herman P. Taub, MediAx Associates

On September 1, 1977 the Administration for Children, Youth and Families (ACYF) of the Department of Health, Education and Welfare awarded a contract to MediAx Associates to develop Head Start measures of program effects on children 3 to 7 years of age. This paper provides an overview of the project and its recommendations.

In carrying out this project, the staff found it necessary to question, reevaluate and recommend major changes in many of the traditional approaches to measurement of program effects on children. I have chosen to prepare this overview from the vantage of these changes, to describe the issues which involved our developing new or different procedures, with full recognition that this is an unsafe strategy for someone, like myself, who does not have a strong background in this field. It is particularly dangerous at a meeting of scholars who might have long and deep attachments to traditional procedures. But I, and my colleagues, have decided to bring out these differences, rather than avoid them, for three reasons:

1. If there is a need for significant change, as we have concluded for so many of the previous approaches, it is important that a dialog on such issues be initiated at the earliest possible time so that we, and ACYF, have the

*Presented at the AERA annual meeting, April 11, 1979.

greatest possible assurance of adopting recommendations based on full and open consideration by the profession.

2. The potential influence of measures of program effects is so great, and affects such critical areas, that it is proper that purported inadequacies in present approaches, and all recommended new approaches, be aired and understood by the broadest possible audience and at the earliest possible time.

3. Whatever the final outcome, the ultimate purpose of this project is to provide guidance to policymakers, to administrators, to program staffs and to parents, on the practices that will ultimately affect the development of a large proportion of young children in the United States. As such, this project must merit, and needs, your support and involvement. We hope that this presentation, given while the project is still in its formative stages, will initiate a dialog which will ultimately involve a significant proportion of those concerned with early child development.

At this point, we need to clarify that although ACYF, and the Department of Health, Education and Welfare, deserve full credit for undertaking and supporting this project, and actively participating whenever requested, they have not yet reviewed or approved many of the positions taken in this paper. Therefore, neither they, nor anyone else, associated with this project, should be assumed responsible for any of the errors, omissions or poor judgements in what I here present.

The first need for change that we encountered on this project was in its name. This was the harbinger of the nature of this entire project. It was originally named "Development of Measures of Social Competence" to provide recognition for the full range of developmental concerns of Head Start and the need for an equally broad system of measures. However, it soon became apparent that use of the term "Social Competence" carried with it two major liabilities which more than offset, in this project, its many advantages.

First, to many in the field of child development the term did, and still does, imply an emphasis on the social development of children rather than on the broad range of concerns of Head Start, and of this project, which include all aspects of children's development that may affect their general overall competence, as children and in later life. Secondly, to others, the term unfairly implied that children in Head Start were particularly incompetent socially and that "Washington" was attempting, through this project, to establish and impose standards of social competence for all children. This latter argument, and the correlant issue that it would be invalid psychometrically and destructive to children to measure program effects in terms of national status norms, was first impressed on us by the parents and staffs of Head Start that we spoke with around the United States. When we brought our concerns on the name back to ACYF, they were most understanding and supportive, as they have been on all of the issues in this project. The new name that was officially adopted is "Development of

Head Start Profiles of Program Effects". The use of the term profiles is intended to communicate a range of characteristics which need to be individually attended; we particularly like the term profiles because it also suggests -- a set of data on many features that, when viewed as a whole, portrays a unique character.

One of the first technical issues in this project involved the selection of characteristics to be included in the measures. It soon became obvious that a valid approach to program effects measurement must encompass the broad range of concerns of parents, program staffs, public school teachers, child development specialists, program administrators and program policy makers. To leave out any area of interest would be permitting those audiences concerned with omitted characteristics only a partial and inherently distorted perception of program effects. It also became apparent that the primary purpose for this project, to all those concerned, could be generally described as providing information of use in the improvement of programs.

To obtain specific descriptions of desired characteristics this project has collected data on desired characteristics from 22 recognized child development scholars representing a wide range of perspectives, from 176 parents, from 134 Head Start Staff, from 64 public school teachers of kindergarten through second grade, from a review of the literature and particularly that of previous Head Start funded projects which

defined characteristics for measurement, and from special studies commissioned under this project in special areas.

The scholars were invited to submit their recommendations, on characteristics to be measured, in response to an open ended question from J. McVicker Hunt. Twenty two, of the 31 scholars invited, responded.

The parents, staff and teachers were selected to be representative of the diversity in the populations served by Head Start. They attended two-day workshops held in New York City, Atlanta, Dallas, San Francisco, Kansas City and Chicago. Overall, 36% of those attending were Black, 19% Hispanic, 5% Native American, 4% Asia and 36% White and others. They represented rural, urban and migrant populations. They met for one day in small homogeneous groups of 4 to 5 to discuss and list specific behaviors that would indicate desired development of children in the various areas; on the second day they individually evaluated the listings of all groups on the first day.

It should be noted that neither the scholars nor the participants to the workshops were selected as a stratified random sample of their respective populations and, therefore, their judgements do not provide a basis for firm generalizations. It is highly probable, however, that the overall range of desired characteristics identified from all these sources closely approximates the range one would find from a statistically representative sample.

Over 1700 different characteristics were identified from these various sources covering all aspects of child development. It was also apparent from this data that there is significant diversity in desired characteristics among scholars, and also among parents, staffs and teachers from different regional and racial-ethnic backgrounds.

The large number of desired characteristics introduced the issue of aggregation. To report program effects in terms of such a large number of individual characteristics would be impractical operationally and, even if attainable, little use could be made of such a "laundry list" of aspects of child development. Upon analysis, all of the desired characteristics could be covered at a higher level of generalization by 26 dimensions.

The 26 dimensions are conceptualized on a still higher level of generalization as coming under six domains. The 6-domain, 26-dimension structure included in this report as Table I, is hypothesized as including all of the characteristics identified through the various sources.

The six domains are Health and Physical Development, Social-Emotional Development, Cognitive Development, Aesthetic Development, Moral and Ethical Development and Life Skills. The characteristics included under Aesthetic Development and under Moral and Ethical Development, and their respective dimensions totalling four in number, are not being recommended for

TABLE I

Six Domains and 26 Dimensions of Child Development

- I. Health and Physical Development
 - 1. Medical
 - 2. Dental
 - 3. Nutrition
 - 4. Physical (Gross Motor)

- II. Cognitive Development
 - 6. Thinking Skills
 - 7. Perception
 - 8. Vocabulary
 - 9. Listening and Speaking
 - 10. Reading and Writing
 - 11. Nature and Science
 - 12. Social Order and Technology
 - 13. Number Concepts

- III. Social Emotional Development
 - 14. Sensitivity of Feelings
 - 15. Expression of Feelings
 - 16. Self Concept
 - 17. Independence
 - 18. Sharing and Competing
 - 19. Peer Relationships
 - 20. Adult Relationships

- *IV. Aesthetic Development
 - *21. Concepts of Quality
 - *22. Concepts of Beauty

- *V. Moral and Ethical Development
 - *23. Concepts of Human Relationships
 - *24. Concepts of Purpose of Life

- VI. Life Skills
 - 21. Interpersonal Competencies
 - 22. Task Competencies

*Not proposed to be included in Head Start Profiles of Program Effects on Children.

measure in this project. They are not specified in the Head Start objectives and, considering the overall needs of measures, it would be difficult to justify expenditures on these areas at this time. We hope that these two important domains of child development will soon be the focus of other projects.

Remaining for this project to consider, therefore, is a four-domain, 22-dimension structure. Each of these four domains, and their respective sets of dimensions and characteristics, has unique qualities aside from the nature of the characteristics being measured.

It should be noted that although we speak of measuring children's development in dimensions and domains, the only direct measurement on children proposed, or possible, is of individual characteristics. We can only develop data on aggregations of characteristics, i.e., on constructs such as dimensions or domains, by statistically combining data obtained on individual characteristics or by obtaining judgments of children's development in the overall construct from persons who are familiar with the particular children and the particular construct. The data obtained on desired characteristics in this project, however, suggests that there is considerable diversity in individual judgments on what constitutes desired or attainable development in children of any age. This need for measures of program effects in aggregated constructs of child development, which inherently depends on diverse judgments, has important implications on the system here proposed.

Desired development in the characteristics and dimensions of the cognitive domain are largely determined by children's experiences and can be represented by a single line without end. There is no conceptual limit on desirable knowledge or skills, assuming no adverse side affects. And we can measure children's development in cognitive dimensions - e.g., thinking skills, reading, number concepts, etc. -- without including characteristics of any other dimension.

In contrast, desired development in the characteristics and dimensions of the social-emotional domain are expressed in terms of balances that are culturally and situationally determined -- e.g., not too sensitive or too insensitive to others' feelings, not overly independent or dependent, not excessively shy or aggressive, not too sharing or too competitive. Also, the dimensions in the social-emotional domain include cognitive attainments -- e.g., knowledge of feelings communicated in verbal and non-verbal expression, realistic knowledge of one's own strengths and limitations, knowledge and skills to function independently, knowledge of customary and expected behavior in peer and adult relationships. Thus, desired development in the dimensions and characteristics of the social-emotional domain are defined by a culturally-determined balance condition. And cognitive characteristics, along with distinctive social-emotional characteristics, are included in the dimensional aggregations of the social-emotional domain.

In the health and physical development domain, however, desired development in particular dimensions and characteristics

is represented by the attainment of a clearly defined status, biologically determined. There is a conceptual ceiling, for example, on desired nutrition or desired fine or gross motor development. And in each of these dimensions we can hypothesize that cognitive characteristics should be included -- e.g., knowledge of health, knowledge of nutrition, dental care skills, how to practice fine and gross motor skills, knowledge of limits of physical undertaking and knowledge of particular sequences involved in motor skills -- are all important to development in their respective dimensions. Also emotional characteristics seem very much involved in defining status of health and physical development -- e.g., health, nutritional and dental attitudes and habits, ability to function in a range of tension-producing environments, self-confidence in undertaking and carrying out motor activities -- all seem part of essential development in their respective dimensions. As such, the dimensions in the domain of health and physical development also have their own distinctive structure - i.e., attainment of a biologically-determined optimum condition - and they include cognitive and social-emotional characteristics in their make-up.

The life skills domain also has a unique structure. The domain is required conceptually to define an integration of development in all of the other domains, to allow for the analyzing, decision-making, motivation and implementation associated with competence in specific situations and to allow for the conceptualizing of alternatives in a given situation and the

selection and implementation of appropriate behaviors appropriate to one's own needs and the needs of the situation. As such, life skills is conceptually of a different and higher order than the other domains. Development in the life skills domain is viewed as being measured in terms of the range and level of complexity of appropriate behaviors - a most challenging but important undertaking for our concerns for child development. And, as indicated, development in the two dimensions of the life skills domain, (1) task competencies and (2) interpersonal competencies, includes characteristics from all the other domains.

Thus, we see that there seems to be an order in our domain structure: cognitive dimensions made up of only cognitive characteristics; social-emotional dimensions including both cognitive and social-emotional characteristics; health and physical development dimensions made up of characteristics of that domain and of the previous two; and finally, life-skills development comprised of its own characteristics and some from all of the other three domains.

Also, there seems to be a unique structure to measurement in each of the domains - cognitive measurement represented by a single, endless line, dependent on experiences; social-emotional represented by a culturally determined balance condition; physical and health measures represented by a biologically determined optimum status; and, life skills measures by the range and complexity of appropriate behavior in one's repertoire.

Apart from the value and implications for this project of insights into the distinctive qualities of the domains, it is reassuring to find these intrinsic differences. Although there may be other domain structures that may be equally usable as the one being recommended here, it is apparent that the four domains proposed for use in this project are not completely arbitrary, in that they seem to reflect a "natural order".

It should be noted that, as pointed out in discussions on the qualities of the domains, the various domains and dimensions are not orthogonal constructs of child development. Rather they are represented as a system of overlapping constructs, each with a generally accepted meaning, that cover most all of the range of desired child development characteristics. As such, they could provide useful first-order information for the dialog among those interested in evaluating and improving child development programs.

However, there are also traits and other constructs, which are frequently used in child development, which are not proposed to be specifically measured in this project at this time. They carry such frequently used labels as curiosity, motivation, locus of control, and many, many more. The particular characteristics they embrace are believed to be included in the characteristics associated with the identified dimensions. It may be found useful at some other time to add or substitute other constructs to the structure here proposed. It is our belief at this time, however, that the identified dimensions meet necessary and sufficient conditions and represent an efficient system for aggregation. Based on a preliminary analysis of the characteristics surveyed, there

appears to be a minimum of overlap in the characteristics included under each dimension, although some is inevitable - e.g., the ability to feed oneself has been included under both "nutrition" and "independence". Such multiply-used characteristics account for less than five percent of the total.

As we have pointed out previously, there has been great diversity in the characteristics identified as being desirable by the various sources. It is here proposed, therefore, to allow for these different values, and for possible different patterns of development among children, in the procedures for selecting characteristics to be measured, for selecting the methods of aggregation, and for validating the measures and aggregation procedures.

These procedures may result in different scales of development being developed and applied, in particular dimensions, with children of different backgrounds. Ratings of children's development in the various dimensions are to be obtained from parents and teachers of children representing the different populations served by the programs, and from appropriate molar measures of development in the various dimensions being evaluated. Characteristics will be selected from the battery directly administered to the same children on which ratings were obtained that yield data, when aggregated appropriately, that most closely correlates with the overall ratings. It is, therefore, conceptualized in this project that one set of measures may be processed differently with children from different backgrounds in order to generate,

where required, a plurality of scales of development in a particular dimension.

We recommend such allowance for differential values being assigned specific characteristics, based on children's backgrounds, as an essential procedure to accommodate the plurality of "good" patterns of child development in the United States and, in particular, those patterns that are most appropriate for children of different cultural backgrounds. Such allowance for different paths of development is consistent with child-development theory, with many philosophies of early childhood, with the policies of many agencies providing child-development services, with the views of the parents, Head Start staffs and public school teachers obtained in this project, with the policies of Head Start, and with the ideals of democracy.

However, the standardized tests that are most available, and therefore most frequently used in program evaluations, are based on a single norm of child development. As such, test items that are sensitive to alternative paths of development would be rejected during test development, if the tests were normed and validated on children of multi-cultural backgrounds. (If they had not been normed and validated on such children the tests would be biased and unfair). It is anticipated therefore that standardized tests will be of limited use to this project in dimensions of development that are culturally affected.

The scales of development that result from the procedures described above will be scales of children's status.

Evaluations of program effects, however, are concerned with change in status rather than attained level of status as the critical variable.

Program effectiveness should be evaluated in terms of the relative change in children's development resulting from the program. Overall program evaluation, by this thesis, should be in terms of the qualitative relationship between relative effectiveness of programs in affecting change in the various dimensions of child development for particular groups of children. Also, it should not be assumed that all program sites would, or should, strive for maximum change in all dimensions of development at the same time; allowances should be made for locally determined differences in objectives.

The implications of these positions prescribe the requirements for proceeding from measures of individual children's status to profiles of program effects on children. First, we are requiring that programs be held accountable for change in children's status, rather than the attained level of children's status. Second we are requiring that there be allowance for different standards of change for different homogeneous groups of children. Third, that these standards be in terms of the maximum change affected by the consistently most effective programs in a particular dimension for particular groups of children. Fourth, that there be allowance for variations among local programs in objectives and emphasis in the development of children in different dimensions. And fifth, that a judgment of overall program effectiveness be based on a qualitative evaluation of the changes affected by the program in all

dimensions of child development. Although each of these requirements has been developed after considerable discussions among the staff and consultants to this project, they are here briefly presented as logical conclusions only because of the need to adhere to the overview purpose of this paper.

The first requirement that programs be held accountable for change in children's status involves obtaining data on children's pre-program and post-program status, and developing respective indices for each in all of the twenty-two dimensions of child development previously described. It is not feasible, or necessary, that data be obtained on all children in all dimensions; a multiple matrix system for selecting children is, however, required for selecting children for particular measures in a manner that will permit maximum analysis of multivariant relationships.

Also, alternative data-gathering techniques need to be investigated to determine which approaches yield reliable and valid data, most cost-effectively, that maximally contribute to the variance in the respective characteristics and dimensions. Furthermore, alternative approaches are required for typical and maximal performance and operant and respondent behaviors, as appropriate, for various characteristics. And, there are a number of characteristics which have been hypothesized as being of considerable significance to measurement of particular dimensions which will require the development of new, high-risk techniques and instruments. For all these reasons, an array of possibly redundant data-gathering approaches,

(with due consideration for cost to potential benefit relationships - potential cost benefits) is being recommended for development, pilot-testing and evaluation during the test development phase of this project, scheduled approximately from Mid-1979 to Mid-1981.

The second requirement, that there be allowance for different standards of change for different children, comes out of the hypothesis that pre-program experiences of children result in a diversity of current developmental potentials, including attitudes, that affect immediately-attainable program change in the various dimensions. Variation in ages of attainment of Piagetian stages of cognitive development is an example of such diversity. The differences in development between children with primarily concrete experiences in rural outdoor settings and those with heavily symbolic experiences involving much language and numbers are hypothesized as another example of children being differentially prepared at a given time for development in the various dimensions.

Thus, it is here hypothesized that there are entry characteristics that affect the rates of children's development during the program experience, and hence the potential for program effects in the respective dimensions, and that these characteristics can be empirically determined through analysis of data on the pre-program and post-program status of large numbers of children. These characteristics have been designated in this

project as the Critical Entry Characteristics (CEC's). To facilitate the identification of CEC's during the field-testing phase of the project, measures have been proposed of attitudinal and other characteristics that are hypothesized as affecting development in specific dimensions - e.g., (1) expectation of success and (2) anticipation of reinforcement. These hypotheses are to be tested, along with the others, through the analysis of data gathered in the field.

Given groups of children with different CEC's, and the change in their development over the program period, then it is possible to identify those programs which consistently affect the largest average gains for each entry group is proposed to be calculated as the average gain of, say, the 10 percent of children with similar CEC's who develop most in the consistently most effective programs. Thus, for any given group of children with similar Critical Entry Characteristics, the effectiveness rating of the program as a whole (i.e., all program sites being evaluated) becomes the percentage that the average gain of all children with similar CEC's is of the average gain of this top 10 percent of children with comparable entry-profiles.

Thus, in using the gains of the consistently most effective programs as the standard, we are creating, at least in the overall, comparable scales of program effectiveness for all children and in all dimensions. Intermediate points on the

scales can be defined in terms of the pre-program distributions of characteristics among children. The overall program-effectiveness rating for a given dimension may then be determined by aggregating (with weighting to reflect the various numbers of children with each rating) the program-effectiveness ratings for the several CEC groups of children involved.

In calculating the effectiveness rating of the program as a whole, allowance needs to be made, as prescribed in the above, for local priorities, since it would obviously be unfair and counterproductive to evaluate programs on the basis of objectives they do not seek to attain. The approach to such allowance being proposed here is to assign relative weights to the several dimensions of child development at a given site on the basis of their relative importance, and then to adjust the gains of children for each dimension at that site according to these weights. After developing comparable indices of program effectiveness for each dimension, it is possible to prepare a profile of overall program effectiveness for each entry group of children in each dimension, and also for all children in all dimensions. Figures 1 thru 4 provide alternative conceptualizations of possible formats for presentation. These formats, and others, will be evaluated during subsequent phases to develop a most effective means of communicating program effects.

It needs to be pointed out, in closing, that the proposed system of measures is not being developed, or recommended, for evaluating individual children or individual programs. As presently conceptualized, it requires large numbers of children to average out errors of measurement and to provide large enough numbers in each of the cells on which calculations are required. It is anticipated that, after the system is in use, stable factors will be identified which will permit reliable use of similar approaches with smaller populations.

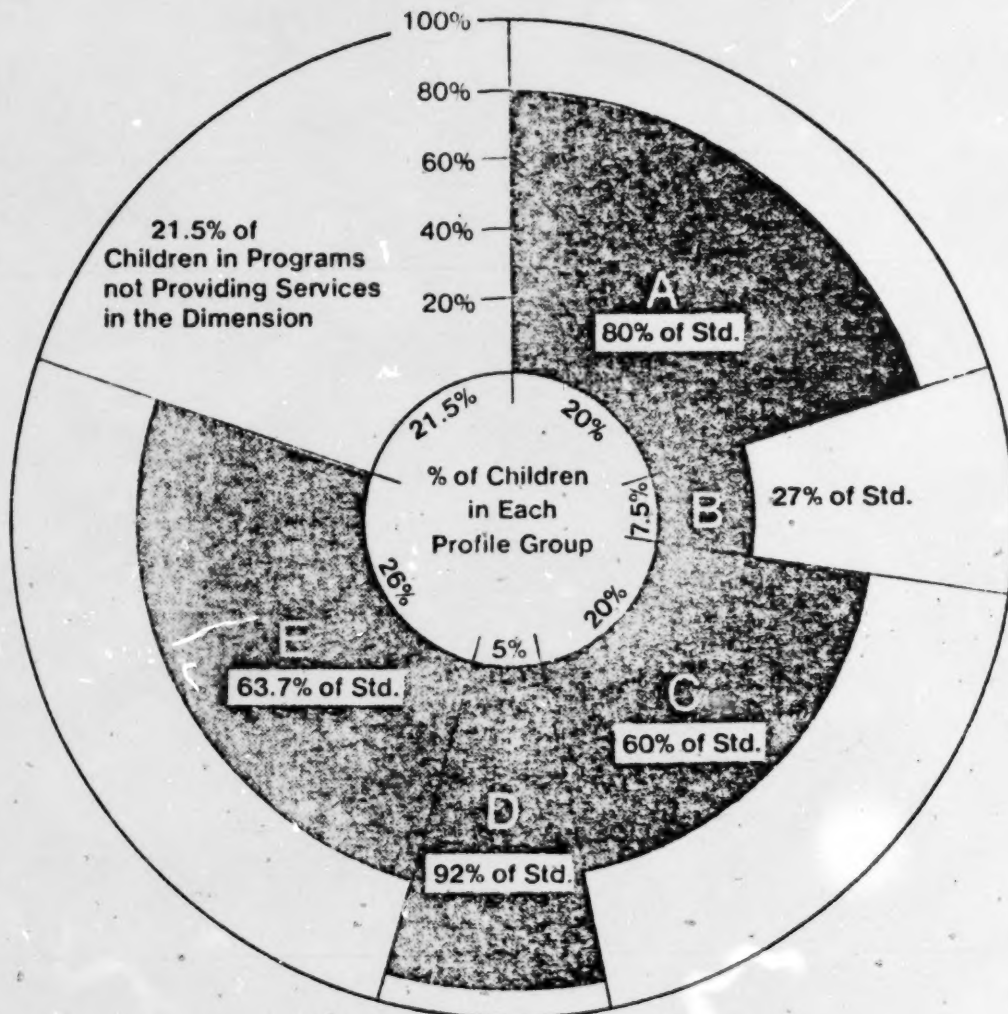
It should also be noted that nothing in the proposed system identifies what program experiences are most effective with children. At best, the system will direct evaluators where to look hopefully to find the program variables that contribute to some sites having the most effective programs.

In this regard, the staff and consultants to the project take every opportunity to point out the need for others to undertake the development of similar approaches to the measurement of program effects on family and community, the other factors greatly influencing child development. With compatible data on program effects on children, on family and on community, it should be possible to gain valuable insights on improving services to young children.

For the present, we are here satisfied with contributing to the impetus that will bring the day when the use of so called

"standardized tests" will be obsolete, along with other well intentioned approaches to providing services that, with the insight of retrospect, we now see did much more harm than good.

Fig. 1 Profile of Program Effects in a Single Dimension
Sample Sector Graph



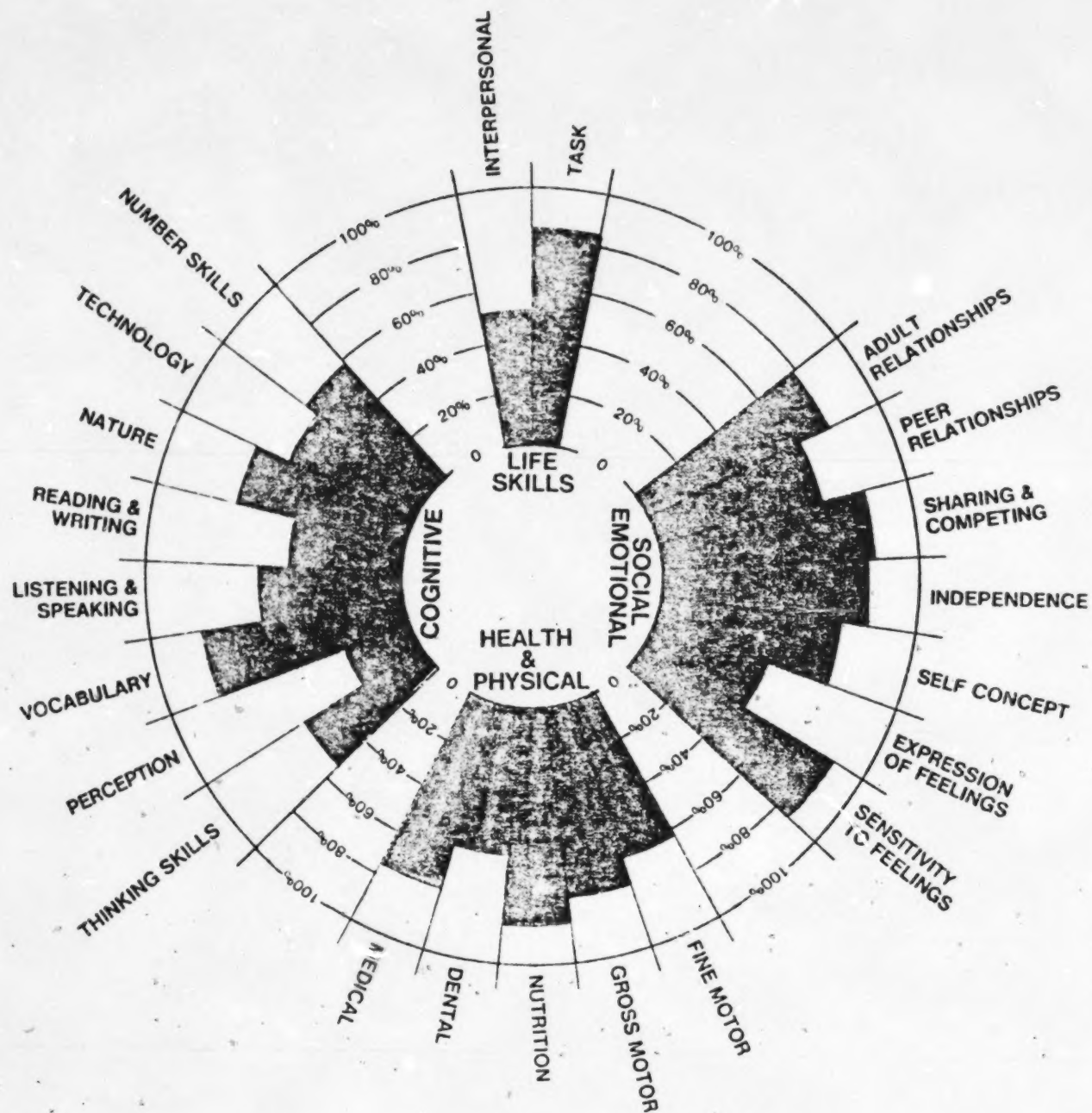


Fig . 2 Profile of Overall Program Effects (Sample Sector Graph)

Fig.3 Profile of Program Effects in a Single Dimension
Sample Columnn Diagram

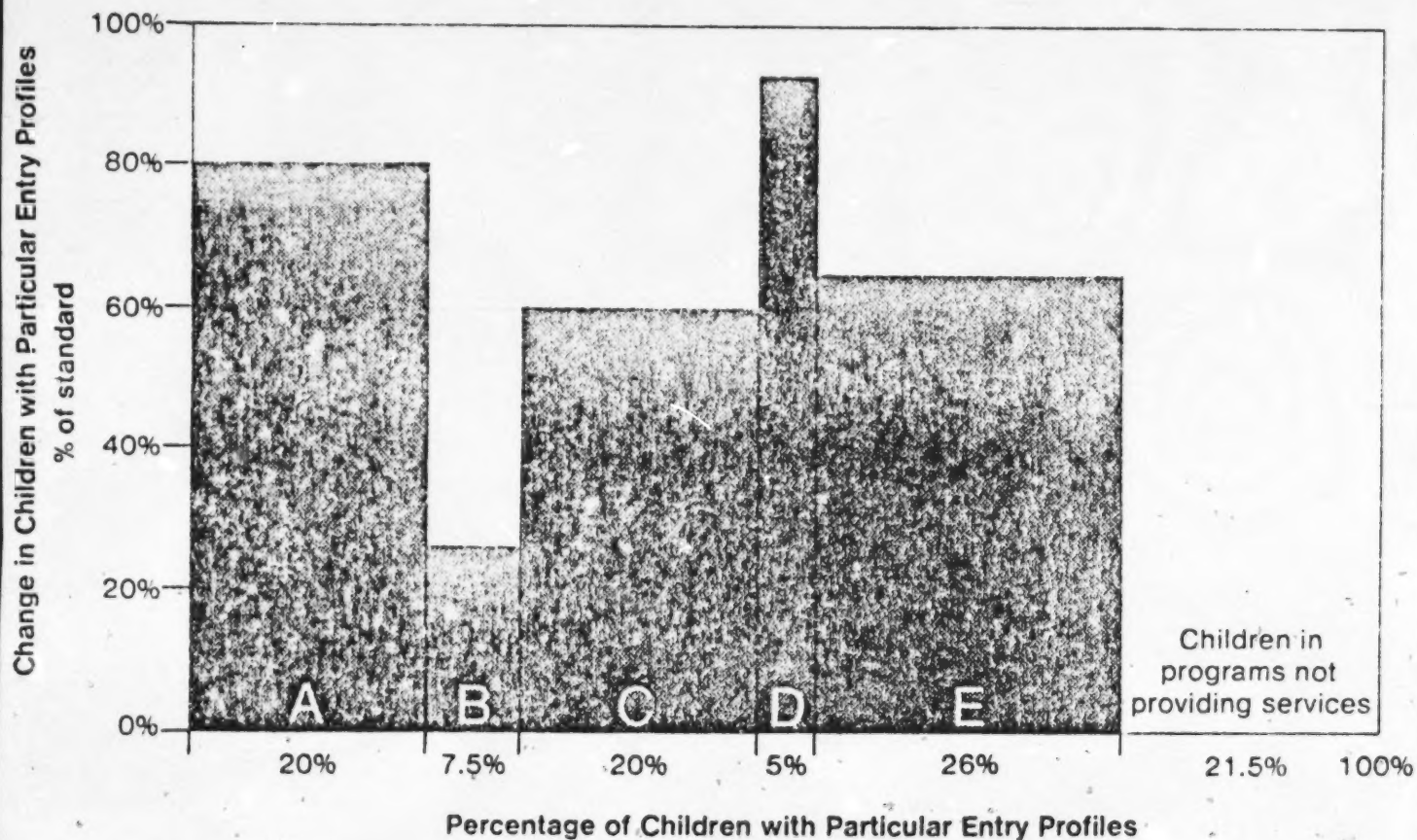
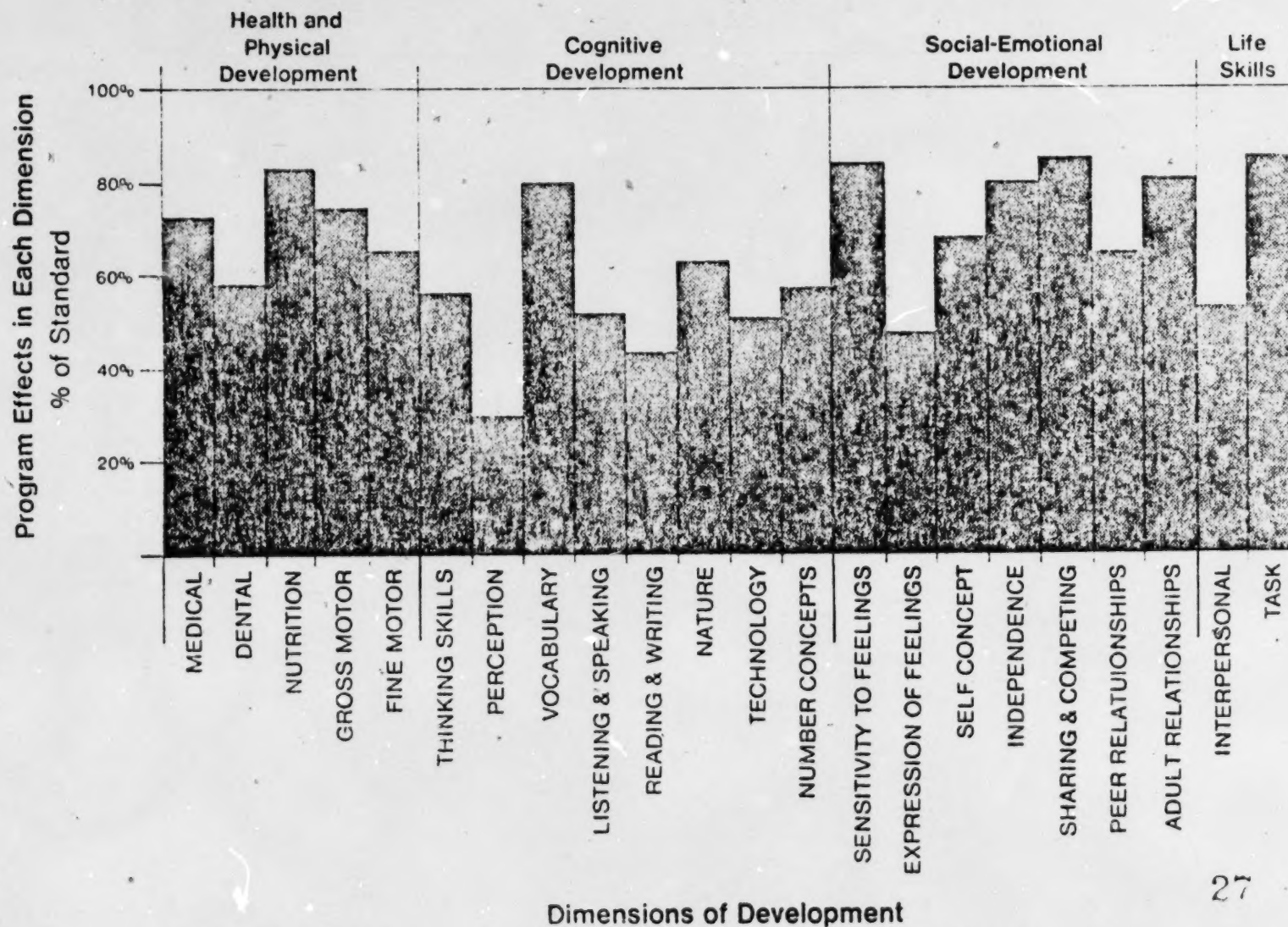


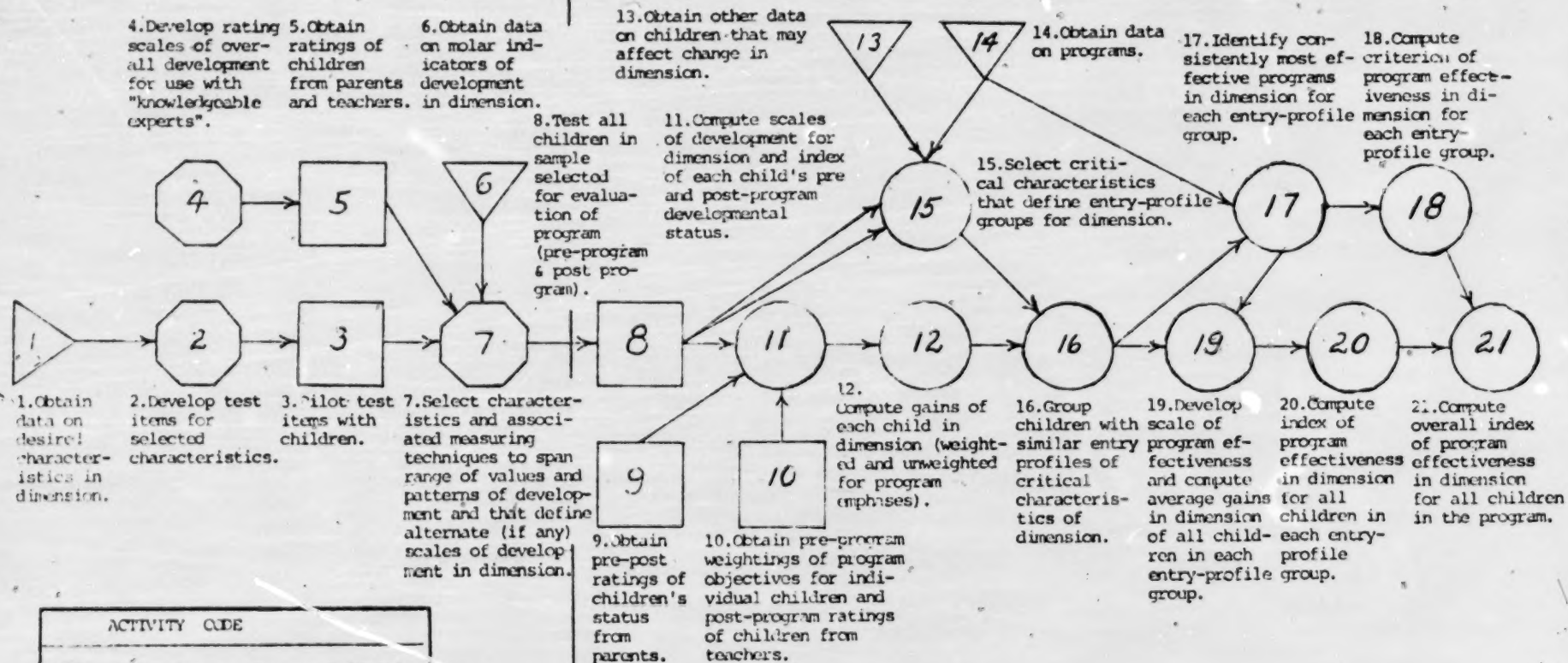
Fig. 4 Profile of Overall Program Effects.
Sample Column Diagram



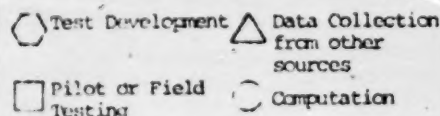
FLOW CHART OF PROCEDURES FOR DEVELOPING & FIELD-TESTING HEAD START PROFILES OF PROGRAM EFFECTS ON CHILDREN

Test Development Phase

Field Testing Phase



ACTIVITY CODE



Note: Development and pilot-testing of statistical techniques and computer program for computations of Steps 11 through 21 to be carried out during test development phase of project.